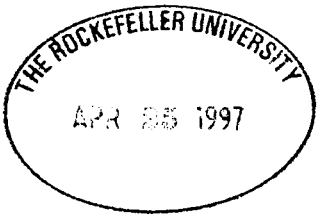
DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

National Institutes of Health
National Library of Medicine
Bethesda MD 20894

April 21, 1997



Dear Panel Member:

Enclosed is a summary of the NLM Long Range Planning Panel on International Programs' second meeting on March 11-12, 1997. We are not making further assignments at this time, although elaborations and comments are always welcome. As we continue to work with the material, and formulate a draft outline of the panel report for your review in June, it is possible that we may get back to you in early May with additional requests.

The meeting summary will be posted on the panel web site shortly. You can find it at http://nlmplan.nlm.nih.gov/international/3summaries.htm.

Drs. Fredrickson Lindberg, Siegel, and I would like to thank you for a most productive and enjoyable meeting.

Sincerely,

Susan Buyer

Susan P. Buyer
Acting Chief, Planning and Analysis
Health Information Programs Development

cc: Meeting Attendees

NLM Long Range Planning Panel on International Programs
Second Panel Meeting
March 11-12, 1997
Meeting Summary

Attendees: see attachment 1

Agenda: see attachment 2

Introduction

The meeting convened in the NLM Board of Regents Room at 8:30 a.m. on Tuesday, March 11 with a call to order by the panel chair, **Dr. Donald Fredrickson.**

Report on NLM Developments since the First Panel Meeting

Dr. Lindberg reported to the panel on developments of interest at NLM since the last meeting. The NLM budget has held up quite well this year. A modest increase is expected for next year; however some congressional leaders intend to support a higher increase.

NLM is dedicated to work on the Next Generation Internet (NGI) to address current problems of security, predictability, and speed. Monies are in the budget for the NGI; the effort is to be managed by the big science agencies (including NIH). Telemedicine will be a major application for the NGI, requiring the health community's participation in its planning and implementation.

Important public policy issues receiving major governmental attention include cloning of mammals and humans.

Dr. Lindberg reported on a recently published guide to assessing telemedicine projects, produced by the Institute of Medicine (IOM) report under contract to NLM. One of the major conclusions of the study was that each telemedicine project should have a business plan that includes evaluation. He also reported on a recent study of the Computer Science and Telecommunications Board and the IOM of the National Academy of Sciences (funded by NLM and the NIH Clinical Center) which examined current privacy and security safeguards for automated patient data in operational health care settings.

The G7 countries met at NLM in January on health care projects, including monitoring the status of the Internet (Dr. Cerf noted that a company named Inverse Technology, http://www.inversenet.com/index.htm, specializes in measuring Internet performance); Visible Human (needs multi language versions); smart cards (the U.S. should catch up as France, Italy, and Germany are strongly committed). As the panel discussed Internet performance, particularly in Europe, Dr. Cerf noted that most UK networks are thinly based re capacity; British Telecom is starting to upgrade. When UK is awake, traffic to U.S. is less, but traffic back is substantial (due to downloading). When U.S. is awake, traffic to UK is less, but return traffic is more. When both UK and U.S. are awake, congestion increases.

There was a discussion of the defeat of the WIPO treaties and discussion of NLM's continuing role in the Presidential High Performance Computing and Communications (HPCC) initiative. NIH is the proxy for DHHS in HPCC; Dr. Shortliffe questioned why DHHS was not represented at the secretarial level. Dr. Lindberg responded that the agencies—such as NIH—have funds to undertake HPCC projects, unlike the Office of the Secretary. Ms. Cade stated telemedicine and health applications must have strong standing within the HPCC effort.

New Content Areas: Molecular Biology Databases

Dr. Lederberg chaired this session, and introduced **Dr. Roberts**, who discussed international aspects of molecular biology databases. A striking example of international cooperation has been the development of the GenBank/EMBL/Database of Japan Consortium of DNA sequence databases. These three databases between them have divided up the task of gathering and incorporating sequence data from their respective constituencies. At the moment this is working well, but the rate of DNA sequence generation and the extent of countries participating in it continues to increase. Further outreach, especially to some of the developing countries, will be required if the databases are to stay up-to-date. The task is made easier by an ethic within the scientific community that new sequences must be deposited in these databases. One challenge that is currently being addressed in the developing countries is cooperation with individual Patent Offices so that sequences deposited with patents are also incorporated into the sequence databases. Further efforts to include lesser-known Patent Offices in developing countries will be required in the future.

Probably the most significant problem for the sequence databases is the inclusion of annotation information, as biological facts about individual sequences become known. This is problematical both at the National and the International level. This is an area where NLM/NCBI could take a lead in devising models for cooperation between scientists and databases and promulgating those models both in the U.S. and abroad.

Other areas of growth at the moment are the so-called mutation databases that contain information about DNA sequence mutations that have been associated with disease. Often the disease-causing genes have been located as a result of family studies, carried out with indigenous populations from countries around the world. Obviously the physicians and researchers in those countries are interested in the outcomes of these studies and have often been key participants in the gene discovery process. There is the opportunity here for NLM to play a leadership role in fostering communications between the databases and the local scientists who helped to develop the information initially. It will be important to maintain a dialog so that the needs of both local physicians and researchers are met as well as the needs of the basic research community.

Just as television has changed the way in which societies communicate over the last fifty years, the Internet is beginning to have a similar effect. Whereas the availability of material over the television was mainly governed by the networks that produce the shows and hence control the medium, the internet is much more egalitarian. In particular, relatively small communities can both develop and share information that

would never be commercially viable. Before too long much of the material that is prepared for the world wide web will be viewable from a home TV set or perhaps more importantly, from a bedside TV set. Here lie some real opportunities for innovation and international collaboration.

In the basic molecular biology area, we are witnessing a philosophical revolution in the conduct of the science of biology. As we obtain complete genome sequences for the organisms around us, the challenge will be to assimilate the available molecular information with the observable function. This will require much greater organization of the data than we have seen today so those programs can compute on the data and calculate models of behavior. The development of the underlying databases will require the participation of many individuals from around the globe if data gathering and information storage are to be efficient. One key role for NLM in this effort would be to promulgate the use of databases to store biological information and to train scientists both in the U.S. and throughout the world in setting up, maintaining and distributing databases of biological information. There is the possibility for a leadership role either in designing or funding the design of innovative databases and then of encouraging their use. In an ideal world one could imagine a cooperation between libraries, databases, journals and authors such that the publication of biological data included both traditional journal publication as well as database submission. The current model by which sequence data gets into GenBank could be extended into other forms of biological data. If the process itself was made easy and the resulting databases were publicly accessible and useful, then such cooperation by the scientists generating the data could probably be assured.

The opportunities in the area of molecular biology databases include the international collection of DNA sequence data and other biological data to be included in databases accessible throughout the world. The international scientific community both generates and uses this data. NLM can play both a leadership role in funding the design of globally accessible databases and in helping to ensure that they are populated. They can foster collaborations between libraries, publisher and authors and can help to train scientists interested in creating new databases. The key to the use of these databases will be the World Wide Web and NLM should use its own database resources as a lure to encourage libraries and scientists in the developing countries to join that World Wide Web. In many instances free access to MEDLINE, GenBank and other NLM resource databases would be a powerful incentive for developing countries to invest in the resources necessary to bring the World Wide Web to their site. Once links are established, scientists in the developing countries could be producers as well as users of database information.

Dr. Lederberg agreed with Dr. Roberts' statements. He noted that there is a tendency to think as a database as something broadcast from a central authority that becomes an opportunity for more dialectic, more commentary. Dr. Lederberg believes that every

database should be accompanied by an open listserve or chat group for further discussion and commentary. Search tools are needed.

Dr. Lipman asserted that the relationship between productivity gains and online journals is unclear. Society changes when we make literature available online but in and of itself, this doesn't lead to greater productivity. Computational methods in science are useful and computing on sequence data yields good predictions and moves science ahead. Computers and networks are on scientists' desktops in industrial countries—that money is already spent—and we can piggyback off that expense and publish a lot of journals online. 75% or more of scientific journals won't be printed in the next few years.

40% of GenBank sequence data originate in Europe and Japan; these data are put into the database as soon as they are received at NCBI. There is no charge for access. The other 60% of data (from the US) are put up on other foreign databases. NCBI serves approximately 20,000 users per day. PubMed is a good example of cooperation with publishers.

In the discussion that followed, Dr. Oettinger stated that this whole area seemed strikingly similar to geographic information systems and related software. Dr. Lindberg agreed and pointed out that the digital library project also is relevant.

According to Dr. Wong, software languages are being developed for specific scientific areas. In the future only a portion of articles will be written in English, the rest in special languages written for gene sequences. Dr. Lipman added that even if we are talking about a system we completely understand, coming up with a robust classification scheme is difficult. If we are talking about a system—nature—that we don't completely understand, the task is daunting. Dr. Cummings informed that panel that a little more than 100 years ago, Dr. Billings engaged a Dr. Burgee for the NLM staff because he was concerned with bacteriology; Dr. Burgee developed a classification system in bacteriology for NLM.

Dr. Lindberg asked Dr. Lipman if there is a problem with slow connections for NCBI databases; there do not seem to be many complaints about connection speed from NCBI users. Dr. Lipman can see from the user statistics when the UK and German business day starts. Fortunately the European and Japanese business days do not overlap, so the traffic load is spread. He does get some complaints but people are patient and he thinks they are willing to try again. However, he pointed out that you can have slower access to the NIH building next door. Dr. Roberts felt that complaints are less because it is still by far much faster than any other means.

Dr. Lindberg brought the group back to Sir George Alleyne's question of what do we need to do now, to get something right away. Somehow we need to separate those needs out.

Dr. Detmer pointed out the need to consider the broader issue of the proprietary nature of information, and, per Ms. Solomon, intellectual property issues. Dr. Elliot wants arguments to deploy against those who want laws such as the WIPO treaties, making databases not free. Ms. Cade is concerned that Congress is looking to get more value from public information without understanding the broader implications for public access.

Preliminary Findings and Recommendations of the Panel

Dr. Fredrickson asked the panel to go around the table and provide one suggestion of a recommendation for the panel report. Subsequent to the meeting, the panel was asked to amplify their comments in writing, and the written responses, where available, have been used in the following:

(in alphabetical order)

13. **Lederberg**—MEDLINE is de facto a filter. Filtering is an important aspect of intermediation. Professional societies should be used to tag what are credible journals. Electronic publications should be embraced. Regional databases can be hot linked to MEDLINE. Anyone can do regional databases that want to, and others who want to can point to them. There are political issue such as how do we know what is good. This is not NLM's role. However, NLM could include tags conveying quality factors. These tags would be done by others, not NLM. There is a risk of NLM trying to over-control and being perceived as political. Other search engines are available. There are challenges in dealing with electronic journals, authentication of electronic publications, and the electronic peer review process.

Meeting Participants

Panel Members

Dr. Donald S. Fredrickson
Panel Chair
Scholar in Residence, National Library of Medicine

Dr. George A. O. Alleyne
Director, Pan American Health Organization

Dr. Floyd J. Bloom
Science Magazine and The Scripps Institute

Dr. David Brandling-Bennett
Deputy Director, Pan American Health Organization

Dra. Gladys Faba Beaumont
Dirrectora, CENIDS, Mexico

Ms. Marilyn Cade
Director, Technology / Infrastructure AT&T

Dr. Martin M. Cummings
Director Emeritus, NLM

Dr. Don E. Detmer
Senior Vice President, University of Virginia

Dr. Richard G.A. Feachem
Senior Adviser, Human Development Department, World Bank

Ms. Jane Bortnick Griffith
Congressional Research Service
(alternate for Dr. James Billington, Librarian of Congress)

Ms. Frances Groen
University Librarian, McGill University

Dr. Robert W. Gwadz
Assistant Chief of Malaria Research, NIAID, NIH

Dr. Joshua Lederberg
Sackler Foundation Scholar, Rockefeller University

Dr. Anthony Oettinger
Director, Program on Information Resources Policy Resources Policy, Harvard University

Dr. Richard Roberts
Research Director, New England Biolabs

Mr. David Russon
Director-General, The British Library

Dr. Edward H. Shortliffe
Professor of Medicine and of Computer Science and Associate Dean for Information Resources and Technology, Stanford University School of Medicine

Dr. H. Kenneth Walker
Professor of Medicine, Emory University School of Medicine

Dr. Eugene Wong
Professor Emeritus, Department of Electrical Engineering and Computer Science
University of California at Berkeley

Board of Regents Liaison

Dr. Marion J. Ball
First Consulting Group (FCG)
Professor, University of Maryland School of Medicine

Ex Officio

Dr. Philip Schambra
Director, Fogarty International Center, NIH

Ad Hoc Consultants

U.S. Government/State Department